



**QUEEN'S  
UNIVERSITY  
BELFAST**

## Deep Learning based Fetal Middle Cerebral Artery Segmentation in Large-scale Ultrasound Images

Wang, S., Hua, Y., Cao, Y., Song, T., Xue, Z., Gong, X., Wang, G., Ma, R., & Guan, H. (2019). Deep Learning based Fetal Middle Cerebral Artery Segmentation in Large-scale Ultrasound Images. In *The IEEE International Conference on Bioinformatics and Biomedicine: Proceedings* (pp. 532-539). Institute of Electrical and Electronics Engineers Inc.. <https://doi.org/10.1109/BIBM.2018.8621510>

### Published in:

The IEEE International Conference on Bioinformatics and Biomedicine: Proceedings

### Document Version:

Peer reviewed version

### Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

### Publisher rights

Copyright 2018 IEEE. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher

### General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

# Deep Learning based Fetal Middle Cerebral Artery Segmentation in Large-scale Ultrasound Images

Shuo Wang\*, Yang Hua<sup>†</sup>, Yunyun Cao<sup>‡</sup>, Tao Song<sup>\*¶</sup>, Zhengui Xue<sup>§</sup>,  
Xiaoping Gong<sup>‡</sup>, Guanjie Wang<sup>‡</sup>, Ruhui Ma\* and Haibing Guan\*

\*Shanghai Key Laboratory of Scalable Computing and Systems  
School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University,  
Shanghai 200240, China

Email: {gavin.wang, songt333, ruhuima, hbguan}@sjtu.edu.cn

<sup>†</sup>EEECES/ECIT, Queen's University Belfast, Belfast BT7 1NN, UK

Email: Y.Hua@qub.ac.uk

<sup>‡</sup>International Peace Maternity and Child Health Hospital,  
School of Medicine, Shanghai Jiao Tong University,  
Shanghai 200030, China

Email: wangbin369@sjtu.edu.cn, ping20050708@sina.com, ypl01905@rjh.com.cn

<sup>§</sup>School of Computing and Mathematics, University of Ulster, Newtownabbey BT37 0QB, UK  
Email: Z.Xue@ulster.ac.uk

<sup>¶</sup>Corresponding author

**Abstract**—Hemodynamic parameters of the middle cerebral artery (MCA) have great clinical value for diagnosing poor pregnancy outcomes and adverse perinatal results. However, in the process of hemodynamic parameters measurement of MCAs from ultrasound images, sonographers have to manually adjust many different parameters for obtaining a clear color flow image, and then adjust the position and size of the gate to obtain an effective spectral image. To reduce the workload of sonographers, we simplify this procedure by introducing a novel deep learning based system, named as MCANet, with which we can directly obtain the desired position of the gate. In order to implement this system, to our best knowledge, we build the first large-scale MCA dataset consisting of 4005 ultrasound images. Considering that the shape and boundary of MCA determine the most effective position of the gate, we firstly segment the MCA region from the raw ultrasound image and then generate the corresponding position of the gate accordingly. We also propose a novel evaluation metric, Affiliation Index, for measuring the effectiveness of the position of the output gate. Extensive experimental results show that our proposed system outperforms the other state-of-the-art methods in terms of Affiliation Index and all other common evaluation metrics on our proposed dataset.

**Index Terms**—middle cerebral artery, ultrasound image, image segmentation, deep learning

## I. INTRODUCTION

The number of women who face a high-risk pregnancy is increasing nowadays due to the delayed marriage and child-bearing. The growth of a fetus can be influenced by pregnancy complications seriously, which are likely to leading poor pregnancy outcomes. Doppler ultrasonic blood flow signals can reflect the changes of blood circulation in a fetus, which is significant to the risk assessment of adverse pregnancy outcomes. Due to its low cost and harmlessness, color Doppler ultrasound

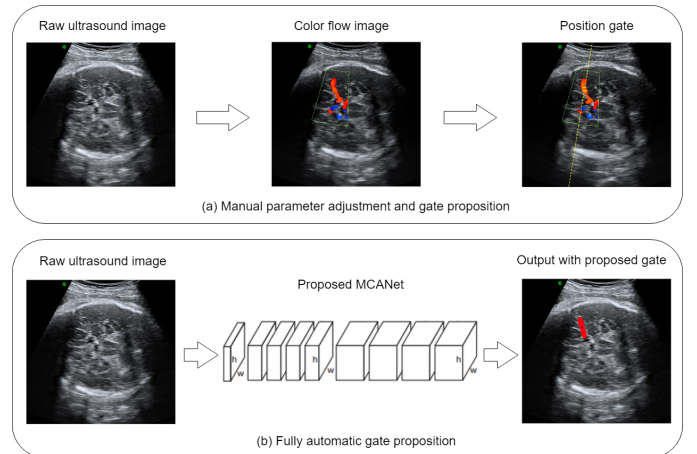


Fig. 1. The comparison between the common procedure of fetal MCA Doppler assessment and the proposed MCANet for the automatic gate proposition. (a) The common procedure of fetal MCA Doppler assessment in which sonographers have to manually adjust parameters for obtaining a clear color flow image and adjust the gate position for obtaining an effective spectral image. In the color flow image, the colored area represents vessels: the red color means Doppler shifts towards the ultrasound beam while the blue color shows Doppler shifts away from it. Gate proposition is the procedure of manually positioning gate for further spectral Doppler imaging. In the last image, the center of the yellow area is usually selected as the gate position; (b) Our proposed MCANet can automatically segment the region of MCA and propose corresponding reliable gate position in the ultrasound image. In the right image, the red color region indicates the MCA and the blue dot represents the proposed position of the gate. *Best viewed in color.*

is commonly used in clinical diagnosis for the inspection of vessels, such as middle cerebral artery (MCA), umbilical artery (UA) and uterine artery to assess fetal development in uterus. Compared with the UA, the MCA is more stable and has less interference factors. Hemodynamic parameters of the MCA have great clinical value for diagnosing poor pregnancy outcomes and adverse perinatal outcomes. However, it is a non-trivial task to obtain the hemodynamic parameters by color Doppler ultrasonography.

Fig. 1 (a) shows the common procedure of the fetal MCA Doppler assessment. Firstly, sonographers need to find fetal standard scan plane for obtaining the anatomical location displayed by an ultrasound image. Then, sonographers try to locate the MCA with the help of the color Doppler flow image (CDFI). Sonographers have to be cautious about adjusting parameters, such as the power, color gain, region of interest and pulse repetition frequency for displaying a color flow image. It is a highly sophisticated task to guide the transducer to obtain the correct scan plane through the highly variable anatomy and to adjust corresponding parameters to obtain a clear color flow image, which requires years of training for a novice sonographer. After obtaining a fine color flow image, sonographers can localize the gate with correct size in order to obtain a spectral image and the corresponding hemodynamic parameters. In real-world scenarios, the gate can be placed at any position in the region of the recognized fetal MCA. For better performance, sonographers commonly set the position of the gate at the center of the recognized fetal MCA region in the color flow image.

In this paper, we propose a novel deep learning based system, entitled as MCANet, which can directly provide the gate position on an ultrasound image with high precision and efficiency. The proposed MCANet can save plenty of time and reduce the workload for sonographers by removing the tedious parameter adjustment in the procedure of fetal MCA Doppler assessment. The workflow of our proposed system is presented in Fig. 1 (b). Note that it is impossible to directly localize the position of the gate due to the extremely small size of the gate area. To address this, we firstly try to detect and localize the relatively bigger fetal MCA region in an ultrasound image and then calculate the center of recognized fetal MCA as the best candidate position for the gate. Because MCA has a variety of shapes, using detection technique for obtaining MCA in a raw ultrasound image cannot offer the corresponding boundary and shape information, which makes it difficult to complete further gate proposition. Therefore, we segment the fetal MCA from the raw ultrasound image and then utilize the center of the segmented MCA as the location of the gate. In order to train and evaluate our proposed MCANet, we collaborate with International Peace Maternity and Child Health Hospital of China welfare institute (IPMCH). We build a large-scale dataset containing 4005 ultrasound images with well-annotated ground truth. Our proposed MCANet can provide real-time inference for the MCA segmentation and the corresponding position of the gate to the sonographers in raw ultrasound image directly. Consequently, it can help simplifying the

operational complexity for inexperienced sonographers and reducing the number of their operation mistakes. Compared with the conventional fetal MCA Doppler assessment system, our system is more accessible to non-experts, which is particularly useful for training purposes. Moreover, our proposed MCANet can accurately predict MCAs in the raw ultrasound images even difficult for experienced clinicians to recognize. This accurate segmentation further helps sonographers obtain the best position for setting the gate. To measure the effectiveness of the proposed gate position by our system, we design a novel evaluation metric, named as Affiliation Index. We also compare our proposed method with the other seven state-of-the-art methods using several evaluation metrics. Experimental results show that our method achieves the best performance with all of these evaluation metrics.

To sum up, our main contributions in paper are as follows:

1. We create a large-scale dataset consisting of 4005 ultrasound images for fetal MCA segmentation and the gate proposition. We will make this fetal MCA ultrasound image dataset publicly accessible.
2. We propose a novel end-to-end training network, MCANet, for completing MCA segmentation and the gate proposition task.
3. We present an evaluation metric, Affiliation Index, for measuring the effectiveness of the gate proposition.
4. Extensive experiments are done by comparing seven different state-of-the-art methods with our proposed MCANet for MCA segmentation and the gate proposition task using Jaccard Index, Dice Coefficient, Hausdorff Distance and our proposed Affiliation Index. Experimental results show that MCANet outperforms others on all of these evaluation metrics.

## II. RELATED WORK

### A. Traditional Methods

Early algorithms of vessel segmentation are based on intensity pattern recognition, e.g., thresholding, which depends on the quantitative differences between a vessel and its surrounding tissues [1]. This kind of method has simple structure and fast speed, but the performance of the algorithms is sensitive to noise. Region growing algorithm is another commonly used method for vessel segmentation. This method starts from some seed points and incrementally recruits pixels according to predefined criteria to segment vessels [2]. It highly depends on the correctness of seed points and can hardly be used in a fully automatic diagnostic system. Some papers use matching filters for vessel segmentation. Different filters are utilized for vessel enhancement in the literature, including Hessian enhancement [3], matched filters [4], Gabor wavelets [5], etc. These matching filters based methods are also sensitive to noise. A single filter generally only works for a vessel.

### B. Machine Learning Based Methods

There are some studies using machine learning methods to obtain noise robust and fully automatic systems for vessel segmentation. In [6], Guo et al. proposed a learning based

vessel detection and segmentation method by using saliency aggregation for detecting vessel candidates, linear filters for pixel level feature extraction and a k-NN classifier for vessel classification. It achieved 56 percent improvement on average precision compared with the hessian based method. However, this method heavily relies on the result of vessel enhancement preprocessing. Moreover, it is time consuming since it starts from detecting vessel candidates and then extracts vessel regions.

### C. Deep Learning Based Methods

1) *Deep Learning Architectures*: Recently, deep learning based methods have been widely employed in medical image segmentation owing to the powerful and promising results. The deep learning network architectures can be broadly grouped into two categories. The first category creates a fully convolutional network (FCN) [7] for feature representation and resolution restoring. This architecture enables itself to learn in an end-to-end model. It is simple and clear but can only make coarse predictions. Although the skip architecture [7] was proposed to solve this problem, the final accuracy of this FCN is still not good enough. The other type uses the architecture consisting of an encoder and a decoder. The encoder is comprised of multiple convolution and down-sampling operations to capture contextual information from raw images. The following decoder learns to upsample decoded feature maps. Moreover, skip connections [7] are commonly applied into this encoder-decoder architecture to combine local and global context. U-Net [8] is a further extension of the FCN which uses decoder to learn upsampling features. However, using decoder to learn upsampling features increases model complexity and computation cost when downsampling is unnecessary needed. Downsampling decreases the resolution of feature maps, which may lose fine details of raw images.

2) *Nature Image Analysis*: Deep Learning based methods have become popular in nature image analysis in recent years. Lin et al. [9] proposed a multi-path refinement network (RefineNet) for semantic segmentation. In their work, they used long-range residual connections to exploit all the information available along the down-sampling process to enable high-resolution prediction. However, the output feature map given by the encoder of RefineNet loses fine structures due to the multiple downsampling. Moreover, by using multi-path refinement, the computation cost increases significantly. Peng et al. [10] presented the method, Large Kernel Matters (LKMs), for addressing both the classification and localization issues for semantic segmentation. They solved the problem by introducing a large kernel. However, the large kernel cannot work well for simultaneously classifying and localizing small objects. In [11], Zhou et al. introduced a dilated convolutional operation on the basis of LinkNet34 [12] which can extract fine structures of small objects. Wang et al. [13] proposed a dense upsampling convolution (DUC) to generate pixel-level prediction. DUC is able to capture and decode more detailed information that is beneficial for restoring small objects but usually missing in bilinear upsampling.

3) *Ultrasound Image Analysis*: Many studies were carried out to analyze ultrasound images based on deep learning. Xu et al. [14] proposed a landmark detection approach for reducing manual efforts and improving the efficiency of clinical procedure for landmark detection. In their work, a supervised action classifier was trained with explicit action steps to generate action maps and successfully translated the landmark detection problem to an image partition problem. Yap et al. [15] proposed the use of deep learning method for automated breast ultrasound lesions detection. They compared three different deep learning methods. Baumgartner et al. [16] developed a CNNs-based SonoNet for real time detection and localization of fetal standard scan planes in ultrasound images. Their designed network can produce high spatial resolution class score maps. By comparing these class score maps with input images in spatial correspondence, they obtained the corresponding localized saliency maps. Chen et al. [17] developed a multi-domain regularized FCN for anatomical structure detection and segmentation from ultrasound images. In their work, they leveraged transfer learning from cross domains to enhance the feature representation. Although their method is efficient in the anatomical structure detection and segmentation, it cannot achieve human expert level performance for many other problems. To the best of our knowledge, there is no work with deep learning for detecting and localizing the MCA in an ultrasound image. Our proposed approach fills in this blank. Compared with the existing deep learning methods applied into other medical images analysis for segmentation, our proposed approach works well on object feature representation, especially for small objects due to higher resolution feature maps. Experimental results show that our method outperforms the state-of-the-art methods in terms of Jaccard Index and Dice Coefficient on our fetal MCA dataset.

4) *Other Medical Image Analysis*: There are large volume of studies based on deep learning for other medical image analysis such as colonoscopy images and X-ray images. Fu [18] utilized the fully CNNs to segment retinal vessels and achieved state-of-the-art performance on DRIVE [19] and STARE [20] datasets. The authors applied FCN-like architecture [21] by adding a side-output layer to fuse different scale information. However, rescaling low resolution output in high layers to the resolution of raw input loses detailed location information, which makes it hard to converge during training. Shvets et al. [22] proposed a deep learning based network with encoder-decoder architecture, AlbuNet-34, for detecting and localizing angiodyplasia lesions and further improved the state-of-the-arts results of MICCAI 2017 Endoscopic Vision SubChallenge: Angiodyplasia Detection and Localization. AlbuNet-34 [22] used a pre-trained ResNet34 network on ImageNet which can benefit from domain knowledge transferring and combining fine location information in shallow layers and semantic information in high layers. However, the pre-trained ResNet34 network is not the best choice for extracting features of small objects since its output resolution is only 1/32 of raw input. All of these works showed the effectiveness of deep learning for medical image analysis.

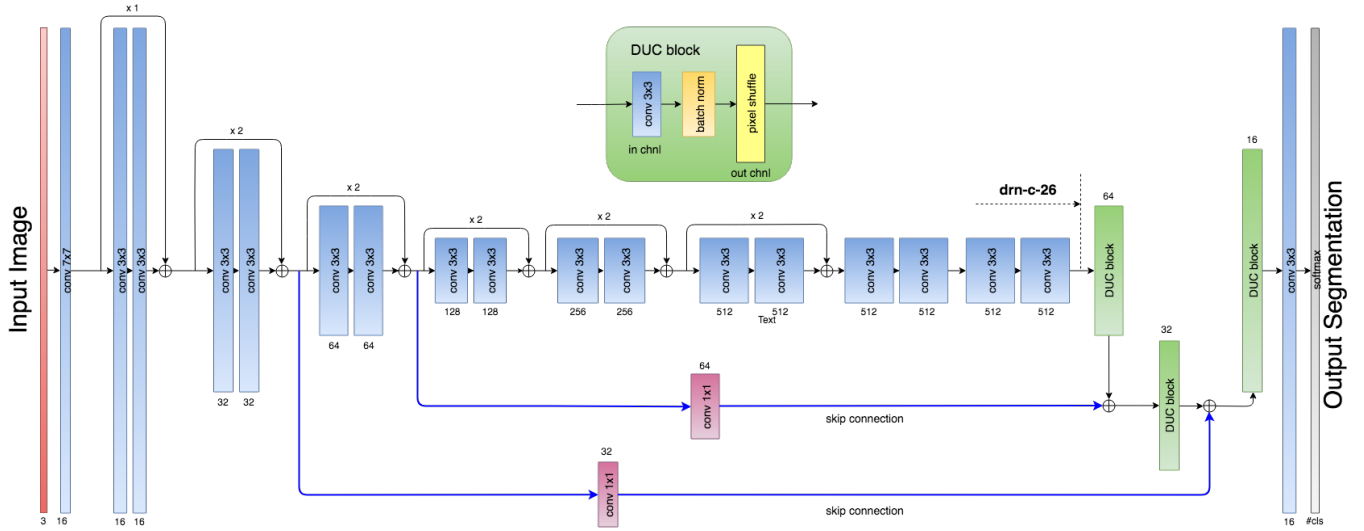


Fig. 2. The overall structure of our proposed MCANet for MCA segmentation in an ultrasound image. MCANet uses pre-trained DRN-C-26 as an encoder and DUC block for resolution restoration. Each box represents a multi-channel feature map and the channel number of each feature map is shown below or above the box. The height of each box denotes the resolution of the feature map. The blue arrows represent skip connections in which local context is transmitted from the shallow layer to the high layer. *Best viewed in color.*

### III. METHODOLOGIES

#### A. MCANet Architecture

Our proposed MCANet for MCA segmentation is built upon the skip architecture with pre-trained dilated residual network (DRN) [23], which is pre-trained on 1000 classes of ImageNet dataset [24]. Furthermore, MCANet also contains DUC part, which is utilized to generate pixel-level prediction. The detailed architecture of MCANet for MCA segmentation is presented in Fig. 2.

Our proposed network starts with a block performing convolution with the kernel size  $7 \times 7$ , stride 1 and padding 3. This initial block is followed by repetitive residual blocks. In residual blocks 2, 3 and 4, the first convolution operation provides downsampling with the kernel size  $1 \times 1$  and stride 2 and the rest convolution operations use stride 1. The residual blocks 5 and 6 use dilated convolution with dilation 2 and 4. The later layers of the network remove residual connections in which dilated convolution with dilated 2 and 1 is used. Removing residual connections in later blocks can successfully eliminate gridding artifacts caused by dilated convolutions, and consequently it improves the accuracy of predictions. Moreover, we use the DUC block rather than the upsampling operation for restoring the decoded feature map. In the DUC block, convolution with  $3 \times 3$  filter module and stride 1 helps with the reduction of dimension and computation. The batch normalization module is helpful for training with large learning rate and the pixel shuffle module works for restoring resolution of decoded feature maps. The skip connections are introduced to combine feature maps in shallow layers of DRN-C-26 and output feature map of DRN-C-26. The only difference is that we add extra  $1 \times 1$  convolution in the skip connections.

1) *Fine Feature Representation and Restoration:* In our proposed network, we use DRN-C-26 pre-trained on ImageNet rather than ResNet-34 as encoder, considering that the area of the MCA in a raw ultrasound image is relatively small. The resolution of the output feature map of ResNet-34 on ultrasound MCA dataset is  $20 \times 14$ , which is only  $1/32$  of the raw ultrasound image. This output is too coarse for decoder in AlbuNet-34 network to restore fine detailed context information. In contrast, DRN gets rid of this problem by removing the max pooling layer. DRN uses convolutional filters for downsampling to remove gridding artifacts caused by the max pooling operation. Moreover, DRN combines dilated convolution to preserve spatial resolution, thereby it produces a comparatively high resolution output feature map. The large resolution feature map contains fine detailed information which is helpful for restoring a small object. Our proposed network takes advantage of the effective skip connections to combine the local information in shallow layers and semantic information in high layers. We also make use of the effective small object restoration ability of DUC to better restore final details in our MCANet.

2) *Computational Efficiency:* We do not add skip connections to all layers in DRN-C-26 to maintain the computational efficiency. Feature map of the last layer has both large resolution and sufficient semantic information, which makes it possible to obtain good performance without adding complex skip connections. Furthermore, DUC module can directly apply convolutional operations on the feature map, which is more computationally efficient than upsampling operations at the same scale.

### B. Loss function

We use a hybrid loss function for optimizing our mode. The hybrid loss is as follows:

$$L = (1 - \lambda)H + \lambda(1 - J), \quad (1)$$

where

$$H = \frac{1}{N} \sum_{i=0}^N \left( y \log y' + (1 - y) \log(1 - y') \right) \quad (2)$$

and

$$J = \frac{1}{N} \sum_{i=0}^N \frac{y_i y'_i}{y_i + y'_i - y_i y'_i} \quad (3)$$

$H$  denotes binary classification loss and  $J$  represents Jaccard loss. In (2) and (3),  $y$  and  $y'$  are the target and prediction vectors, respectively. The probabilities of correctly classifying each pixel, and the intersection between mask and corresponding predictions can be simultaneously maximized by this hybrid loss.  $\lambda$  is the weight for balancing the influence of binary classification loss and Jaccard loss to the final predictions.

### C. Post Processing

The output of the proposed network is binary segmentation mask in which the region of the MCA is labeled as 1 and the background is labelled as 0. In some rare cases, due to the similar components and noise in raw ultrasound images, the output masks of MCANet may contain multiple MCA regions with a variety of sizes. We just simply keep the segmented MCA with the largest area size as the initial prediction.

Moreover, we also apply Gaussian filter to smooth boundaries of the initial predictions, which may contain jagged shape and pointed projections, in order to generate the final MCA segmentation results. Unlike other methods, traditional boundary refinement is not accessible in our framework since there is too much noise in raw ultrasound images to recognize the contour of the MCA.

## IV. EXPERIMENTS

### A. Dataset

A well-defined dataset, especially large-scale dataset, is essential to deep learning based techniques. We build a large-scale MCA dataset for research purpose, collaborating with the IPMCH, which is one of the largest specialized hospitals in China providing the world-learning maternal and child health care services. During the process of constructing this large-scale dataset, we have collected, cleaned and annotated MCA ultrasound images spanning more than half year. Firstly, we collect a total of 5230 raw ultrasound images from 1783 fetuses with gestational ages ranging from 28 to 40. Specifically, these images are acquired with ultrasound systems of Voluson E8 (GE) at the resolution of  $640 \times 448$  with three channels. To obtain a valid raw ultrasound image, sonographers have to be cautious about guiding the transducer to get standard scan plane with only one MCA during an

examination. This requirement increases the difficulties of raw image collection. It typically takes tens of minutes for a professional sonographer to complete the procedure for one fetal MCA Doppler assessment and raw data collection. Secondly, after completing the collection of the raw ultrasound images, we remove the raw images with multiple MCAs or low quality, e.g., containing too much noise or bad color flow. We select 4005 qualified raw ultrasound images out of the 5230 images to construct the ultrasound image dataset. Finally, we have experienced sonographers annotating these 4005 ultrasound images. Using our self-developed annotation tools, sonographers manually annotate these ultrasound images with the help of the corresponding color Doppler flow images. Recognizing and locating MCAs in raw ultrasound images are sophisticated tasks, even for experienced clinicians. The annotated dataset is further reviewed by senior sonographers to make sure the correctness of these annotations.

To our best knowledge, the proposed MCA ultrasound image dataset is the first dataset with sonographer-annotated ground truth for fetal MCA segmentation and the gate proposition. We will make our MCA dataset publicly available for research purpose.

### B. Evaluation Metrics

In order to compare with the other state-of-the-art methods fairly, we apply Jaccard Index, Dice Coefficient and Hausdorff Distance as the evaluation metrics to show the performance of each model in different aspects. Detailed information of the Jaccard Index, Dice Coefficient and Hausdorff distance are shown in (4), (5) and (6), respectively.

- **Jaccard Index** (Jaccard Similarity Index) - measures the similarity between finite sample sets. It compares the prediction ( $A$ ) with the ground truth ( $B$ ). It is defined as the size of the intersection of  $A$  and  $B$  divided by the size of the union, as follows

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \quad (4)$$

- **Dice Coefficient** - gives the pixel-wise similarity between the predicted segmentation ( $A$ ) and the ground truth ( $B$ ). It is defined as

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|}. \quad (5)$$

- **Hausdorff Distance** - measures the shape similarity of the segmented object ( $A$ ) and the corresponding ground truth ( $B$ ). It is defined as

$$H(A, B) = \min_{a \in A} \left\{ \max_{b \in B} \{d(a, b)\} \right\}, \quad (6)$$

where  $a$  and  $b$  are points of sets  $A$  and  $B$ , respectively, and  $d(a, b)$  is any metric between these points. In this paper, we use the euclidean metric for Hausdorff Distance measurement. The unit of calculation result is in pixels.

Even though Jaccard Index, Dice Coefficient and Hausdorff Distance are three commonly used evaluation metrics for

TABLE I  
THE SEGMENTATION RESULTS OF SEVERAL DIFFERENT STATE-OF-THE-ART METHODS.

Method	JC Index $\uparrow$	Dice Coefficient $\uparrow$	Hausdorff Distance (pixel) $\downarrow$	Affiliation Index $\uparrow$	Inference Time (ms) $\downarrow$
FCN8s [7]	0.579	0.723	31.964	0.850	27.746
U-Net [8]	0.601	0.740	30.383	0.854	12.123
TernausNet-16 [25]	0.610	0.748	28.908	0.863	18.739
RefineNet [9]	0.621	0.759	28.181	0.874	29.777
LKMs [10]	0.624	0.762	27.315	0.881	18.614
AlbuNet-34 [22]	0.622	0.760	27.489	0.881	<b>10.805</b>
D-LinkNet [11]	0.624	0.762	26.930	0.884	17.803
MCANet	<b>0.633</b>	<b>0.768</b>	<b>26.402</b>	<b>0.887</b>	16.295
MCANet + PP	<b>0.637</b>	<b>0.772</b>	<b>26.273</b>	<b>0.888</b>	16.545

binary segmentation tasks, they are not completely suitable for the task of gate proposition in an MCA ultrasound image. To obtain a fine spectral image, the size and position of the gate should be correct. Sonographers are inclined to set the gate in the center of the MCA displayed on a color flow image. Therefore, we propose the following metric, named as Affiliation Index, to evaluate the effectiveness of the proposed gate by calculating the distance between predictions and corresponding masks.

- **Affiliation Index** - gives the effectiveness measurement of the location of a predicted gate. It is defined as

$$r = \frac{1}{N} \sum_{i=1}^N ((1 - \alpha)f(A_i, B_i) + \alpha g(A_i, B_i)), \quad (7)$$

where  $N$  is the total number of raw images in an ultrasound MCA dataset,  $A_i$  and  $B_i$  respectively represent the  $i^{th}$  mask and prediction,  $\alpha$  is a weighting parameter on  $f(A_i, B_i)$  and  $g(A_i, B_i)$ . For simplicity,  $\alpha$  is set to 0.5 in all experiments.  $f(A_i, B_i)$  is designed to measure whether the distance between the center of prediction  $B_i$  and corresponding mask is less than a threshold  $\beta$ , as follows

$$f(A_i, B_i) = \begin{cases} 0 & \text{dist}(A_i, B_i) > \beta \\ 1 & \text{dist}(A_i, B_i) \leq \beta \end{cases}, \quad (8)$$

where  $\text{dist}(A_i, B_i)$  is distance between the center of mask  $A_i$  and the center of the corresponding prediction  $B_i$ . It has the following form

$$\text{dist}(A_i, B_i) = \sqrt{(c_{A_{ix}} - c_{B_{ix}})^2 + (c_{A_{iy}} - c_{B_{iy}})^2} \quad (9)$$

with  $c_{A_{ix}}, c_{A_{iy}}, c_{B_{ix}}, c_{B_{iy}}$  denoting the coordinates of the center of mask  $A_i$  and the prediction  $B_i$ , respectively.  $g(A_i, B_i)$  judges whether the center of the prediction  $B_i$  is in the area of its corresponding mask  $A_i$ . It is defined as

$$g(A_i, B_i) = \begin{cases} 0 & c_{B_{ix}}, c_{B_{iy}} \notin C_{A_i} \\ 1 & c_{B_{ix}}, c_{B_{iy}} \in C_{A_i} \end{cases} \quad (10)$$

with  $C_{A_i}$  representing the area inside the contour of mask  $A_i$ .

Moreover, we measure the inference time for a single MCA ultrasound image since our MCA segmentation task is time-sensitive in the real-world scenario.

TABLE II  
COMPARISON OF THE PERFORMANCE WITH DIFFERENT  $\beta$  IN AFFILIATION INDEX.

Method \ $\beta$	5	8	10	12	15
FCN8s [7]	0.649	0.779	0.850	0.888	0.929
UNet [8]	0.658	0.787	0.854	0.892	0.933
TernausNet-16 [25]	0.667	0.798	0.863	0.899	0.939
RefineNet [9]	0.683	0.811	0.874	0.914	0.953
LKMs [10]	0.688	0.821	0.881	0.919	<b>0.956</b>
AlbuNet-34 [22]	0.685	0.811	0.881	0.916	0.952
D-LinkNet [11]	0.693	0.823	0.884	<b>0.924</b>	<b>0.956</b>
MCANet	<b>0.697</b>	<b>0.828</b>	<b>0.887</b>	<b>0.924</b>	<b>0.956</b>
MCANet+ PP	<b>0.698</b>	<b>0.828</b>	<b>0.888</b>	<b>0.924</b>	<b>0.956</b>

- **Inference Time** - measures the time of one forward pass for one image on a Nvidia Geforce GTX 1080Ti GPU.

### C. Implementation Details

We use open source library Pytorch [26] to implement all networks used in this paper. The code for post processing algorithms is developed in Python. The training is done on a 16GB RAM Linux machine with NVIDIA GTX 1080Ti graphics cards.

1) *Preprocessing and Data Augmentation*: Input images have been normalized by subtracting the mean of images in ImageNet dataset and dividing the corresponding pixel standard deviation before being fed into MCANet, because the DRN-C-26 network used as encoder in MCANet is pre-trained on ImageNet. Moreover, for the purpose of alleviating overfitting, we augment our dataset by rotating training images by  $90^\circ$  and flipping them vertically and horizontally.

2) *Training and Validation*: We use Adam optimizer [27] to train our network with initial learning rate 0.0001 for 10 epochs. Our proposed network converges fast because the pre-trained DRN-C-26 can be transferred to this specific MCA segmentation task easily and normalization operation in later DUC block. During training, the batch size is 16 and a 3-fold cross validation is used.

### D. Experimental Results

1) *Performance comparison with state-of-the-art methods*: For complete comparison, we evaluate the performance of seven state-of-the-art methods on our proposed MCA ultrasound dataset, including FCN [7], U-Net [8], TernausNet-16



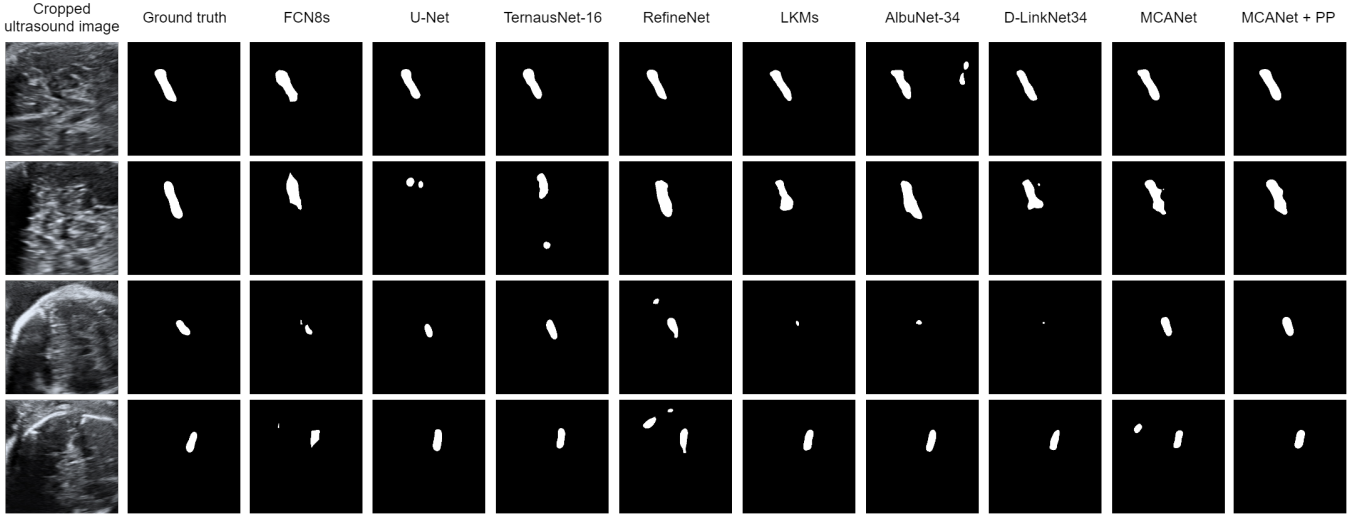


Fig. 3. Visual segmentation results of different models. These input images in the first column are cropped from raw ultrasound images for displaying detailed context of MCA and corresponding predictions (PP denotes post processing).

TABLE III  
IMPACT OF DIFFERENT WEIGHTS IN LOSS FUNCTION ON PERFORMANCES OF FINAL PREDICTIONS.

$\lambda$	JC Index $\uparrow$	Dice Coefficient $\uparrow$	Hausdorff Distance (pixel) $\downarrow$	Affiliation Index $\uparrow$
0	0.604	0.744	31.150	0.860
0.1	0.629	0.765	27.457	0.885
0.2	0.629	0.765	27.359	0.885
0.4	<b>0.633</b>	<b>0.768</b>	<b>26.402</b>	<b>0.887</b>
0.8	0.630	0.766	27.076	0.886
1	0.626	0.763	28.688	0.886

[25], RefineNet [9], AlbuNet-34 [22], LKMs (Large Kernel Matters) [10] and D-LinkNet [11]. U-Net and AlbuNet-34 are two state-of-the-art deep learning models designed for biomedical image segmentation. FCN is the first fully convolutional network for semantic segmentation. TernausNet-16 and D-LinkNet are two winner solutions for Carvana Image Masking Challenge 2017 and CVPR DeepGlobe 2018 Road Extraction Challenge, respectively. RefineNet and LKMs are also for semantic segmentation and both methods achieve promising results on PASCAL VOC Challenge 2012.

The overall evaluation results of all methods are shown in TABLE I. Our proposed MCANet achieves the best results on all evaluation metrics. Specifically, we obtain the Jaccard Index 0.633, Dice Coefficient 0.768, Hausdorff Distance 26.402 and Affiliation Index 0.887 (by setting  $\alpha$  to 0.5). For speed comparison, AlbuNet-34 is the fastest model with inference time 10.8 milliseconds which is almost three times faster than the slowest method, because of its light encoder. MCANet is the third fastest approach for processing raw ultrasound images and its inference time is 16.3 milliseconds. Comparing to AlbuNet-34, MCANet is with more dilated convolution operations, which has better representation power on small objects but takes more memories and calculation time.

Moreover, we show the segmentation results of different state-of-the-art methods in Fig. 3. Obviously, these seven state-of-the-art methods can work well for simple images,

e.g., the ultrasound image shown in the first row. For a more complicated raw input image in the second row, the heavier networks, RefineNet, AlbuNet-34 and MCANet give better predictions, due to their powerful ability for feature representation. According to the third row, for a small MCA in a raw image, the predictions of TernausNet-16 and MCANet are promising. TernausNet-16 uses VGG16 [28] as the encoder and MCANet uses DRN-C-26 as the encoder which both generate a relatively large output feature map. Larger feature map is necessary for recovering fine details of small objects. Comparisons in the first row and last row show the necessity of post processing for removing invalid predictions and smoothing the boundary of final predictions.

#### E. Ablation study

1) *System Performance with Different  $\beta$  in Affiliation Index:* We compare the system performance with different  $\beta$  for our proposed Affiliation Index. The results are given in TABLE II. It is clear to see that when the  $\beta$  value increases, i.e., the tolerance for the error of the center point in MCA segmentation result also increases, the performance of all methods become better and saturated. According to the feedback of professional sonographers, we set  $\beta = 10$ , which balances the tolerance of the error and reasonable performance for the real-world application, and report the results in TABLE I and III in terms of Affiliation Index.



2) *System Performance with Different  $\lambda$  in Loss Function:* The value of  $\lambda$  in loss function, i.e., (1), can significantly influence the final accuracy. In TABLE III, when  $\lambda = 0$ , the hybrid loss function is equal to the binary cross entropy loss. Training with this binary cross entropy loss, we achieve worse Jaccard Index, Dice Coefficient, Hausdorff Distance and Affiliation Index than hybrid loss. When  $\lambda = 1$ , the hybrid loss becomes Jaccard loss, which gives sub-optimal results. Very close results are obtained when the value of  $\lambda$  is set to 0.1, 0.2 and 0.8. When  $\lambda = 0.4$ , we obtain the best performance.

## V. CONCLUSION

In this paper, we developed a computer aided system, named MCANet, for automatically generating desired position of the gate in the process of fetal MCA Doppler assessment. By using our proposed system, sonographers can get rid of tedious parameters adjustment for obtaining a color Doppler flow image and corresponding position of the gate, which will save plenty of time and reduce the number of operation mistakes for sonographers. To implement this system, we built a large-scale dataset containing 4005 ultrasound images and proposed a novel evaluation metric, Affiliation Index, for measuring the effectiveness of the proposed gate position. Comparative experiments show that MCANet outperforms the other state-of-the-art methods in terms of Affiliation Index and all other evaluation metrics used for validating performance in different aspects.

## ACKNOWLEDGMENT

This study was supported in part by National NSF of China (NO.61525204, 61732010, 61872234) and under the interdisciplinary program of Shanghai Jiao Tong University YG2015QN32.

## REFERENCES

- [1] N. Niki, Y. Kawata, H. Satoh, and T. Kumazaki, "3d imaging of blood vessels using x-ray rotational angiographic system," in *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.* IEEE, 1993.
- [2] A. Bhuiyan, B. Nath, and J. J. Chua, "An adaptive region growing segmentation for blood vessel detection from retinal images," in *VISAPP (1)*, 2007.
- [3] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 1998.
- [4] A. A.-H. A.-R. Youssif, A. Z. Ghalwash, and A. A. S. A.-R. Ghoneim, "Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter," *IEEE Transactions on Medical Imaging*, vol. 27, no. 1, 2008.
- [5] Z. Hu, M. Niemeijer, M. D. Abramoff, K. Lee, and M. K. Garvin, "Automated segmentation of 3-d spectral oct retinal blood vessels by neural canal opening false positive suppression," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2010.
- [6] P. Guo, Q. Wang, X. Wang, Z. Hao, K. Xu, H. Ren, J. B. Kim, and Y. Hwang, "Robust vessel detection and segmentation in ultrasound images by a data-driven approach," in *Medical Imaging 2014: Image Processing*, vol. 9034. International Society for Optics and Photonics, 2014.
- [7] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015.
- [8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention.* Springer, 2015.
- [9] G. Lin, A. Milan, C. Shen, and I. D. Reid, "Refinenet: Multi-path refinement networks for high-resolution semantic segmentation," in *Cvpr*, vol. 1, no. 2, 2017.
- [10] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel mattersimprove semantic segmentation by global convolutional network," in *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on.* IEEE, 2017.
- [11] L. Zhou, C. Zhang, and M. Wu, "D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018.
- [12] A. Shvets, A. Rakhlin, A. A. Kalinin, and V. Iglovikov, "Automatic instrument segmentation in robot-assisted surgery using deep learning," *arXiv preprint arXiv:1803.01207*, 2018.
- [13] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, and G. Cottrell, "Understanding convolution for semantic segmentation," *arXiv preprint arXiv:1702.08502*, 2017.
- [14] Z. Xu, Q. Huang, J. Park, M. Chen, D. Xu, D. Yang, D. Liu, and S. K. Zhou, "Supervised action classifier: Approaching landmark detection as image partitioning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2017.
- [15] M. H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwigglelaar, A. K. Davison, and R. Martí, "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE journal of biomedical and health informatics*, vol. 22, no. 4, 2018.
- [16] C. F. Baumgartner, K. Kamnitsas, J. Matthew, T. P. Fletcher, S. Smith, L. M. Koch, B. Kainz, and D. Rueckert, "Sononet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound," *IEEE transactions on medical imaging*, vol. 36, no. 11, 2017.
- [17] H. Chen, Y. Zheng, J.-H. Park, P.-A. Heng, and S. K. Zhou, "Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2016.
- [18] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on.* IEEE, 2016.
- [19] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, 2004.
- [20] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical imaging*, vol. 19, no. 3, 2000.
- [21] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proceedings of the IEEE international conference on computer vision*, 2015.
- [22] A. Shvets, V. Iglovikov, A. Rakhlin, and A. A. Kalinin, "Angiodysplasia detection and localization using deep convolutional neural networks," *arXiv preprint arXiv:1804.08024*, 2018.
- [23] F. Yu, V. Koltun, and T. A. Funkhouser, "Dilated residual networks," in *CVPR*, vol. 2, 2017.
- [24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012.
- [25] V. Iglovikov and A. Shvets, "Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation," *arXiv preprint arXiv:1801.05746*, 2018.
- [26] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *NIPS-W*, 2017.
- [27] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [28] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.